International Journal of Advances in Engineering Sciences Vol 8, No 2 (2018)

EFFICIENT SCALE INVARIENT FEATURE BASED METHOD FOR CROWD LOCALIZATION

Anshit[1], Dr. Pawan Kumar[2], Dr. Deepak Goyal[3], Dr. Pankaj Gupta[4], Monika Goyal[5]

[1] M.Tech Student, Department of Electronics and Communication Engineering, Vaish College of Engineering, Rohtak

[2]Assistant Professor, Department of Electronics and Communication Engineering, Vaish College of Engineering, Rohtak

[3]Associate Professor, Department of Computer Science Engineering, Vaish College of Engineering, Rohtak

[4]Professor, Department of Computer Science Engineering, Vaish College of Engineering, Rohtak

[5] Lecturer– Vaish Mahila Mahavidyalaya, Rohtak (Haryana), India

Abstract

Visual surveillance has been a very active research topic in the last few decade due to growing importance for security in the public areas. With the increasing number of CCTV networks in public areas, the enhancement in the computing power of modern computers and increase the possibility to entrust an automatic system with the security and the monitoring of events involving large crowds is within reach. Crowd detection and localization in the surveillance video is the first step in automatic crowd monitoring system. The performance of the whole system depends on this step. Detecting the crowd is a challenging task because the crowds come in different shape, size and color, against cluttered background and varying illumination conditions. As the size of the crowd increases managing the crowd becomes more complex. Here we introduce a new method for detecting and localizing the crowd in a video scene is one of themost challenging task. This thesis address the problem of detecting crowd from the video captured through the surveillance camera The algorithm needs less resource in terms of time and memory, so it can be used in the real time video surveillance system. The Results clearly indicate the quality of the crowd localization that we have used to track people in the moving scene the algorithm can efficiently work over any given input scene.

Index Terms: Computer Vision, Crowd Localization, Visual Surveillance.

I. Introduction

Human vision system is capable to perform computationally complicated tasks such as detecting or counting similar object in a scene in spite of occlusion and clutter, without many difficulties [1]. Research scientist in the computer vision community has been developing mathematical tools to detect objectsrecognize objects and actions discover behavior and events in visual scenes comparable to human capabilities. In all these efforts understanding of human activity is of aspecial interest for both application and research purposes.

(a) a stage of the tour de france (structured scene)

(b) people gathered in a square (unstructured scene)

Fig. 1. Examples of crowded scenes:

One of the main application areas of computer vision is surveillance system. Video surveillance is the area of computer science devoted to real-time acquisition, processing, and management of videos coming from cameras installed in public and private areas, in order to automatically understand events happening at the monitored sites, eventually sending up an alarm [2]. Because of the rapidly increasing number of surveillance cameras, it has become a key technology for security and safety, with applications ranging from the fight against terrorism and crime, to private and public safety (e.g., in private buildings, transport networks, town centers, schools, and hospitals), and to the efficient management of transport networks and public facilities (e.g., traffic lights and railroad crossings).Large political rallies, ceremonies, and rituals which involve large groups of people pose significant challenges for officials in terms of security, monitoring, and organization. First and foremost, in large gatherings the security of the event is of the highest importance. In dense crowds, any abnormal behavior or incident would lead to a cascade of undesirable events because of the synergic effect of human interactions. Moreover, the larger the scale of the crowd becomes, the harder the visual surveillance is for the human eye. Therefore, authorities advocate the use of automatic tools to analyze crowd behavior and to assist the detection and localization of abnormal events within crowds [3-10].

I.1Proposed Work

Crowd localization is one of the main important phases in developing the automated surveillance system used for the monitoring dense crowded area. Success of crowd localization depends on the segmentation algorithm used for crowd detection [3]. Since the surveillance system works in real time the segmentation algorithm must be fast and require less memory. In this thesis I proposed a new approach for crowd detection and then estimate the size of the crowd successfully.

II. Problem Statement

Detecting and localizing the crowd in a video scene is one of the most challenging task for the researchers working in the field of developing smart surveillance system. This thesis address the problem of detecting crowd from the video captured through the surveillance camera.

II.1Introduction

Proposed crowd localization method consists of following stages:

• Background Estimation

• Segmentation

• Crowd Estimation [4].

II.2 Background Estimation

There are various methods proposed for background estimation:-

These method are broadly classified into two groups:

1. Background Subtraction
2. Optical Flow Analysis.

Algo. 1 Background Estimation Algorithm

1 For each frame $f(t) of the video$,extract Red, Green and Blue channel separately.

2 For each channel i calculate background using the formula

FDBG(x,y,t)=(1-$\propto$)*$fi(x,y,t)$+$\propto$*FDBG$_i$(x,y,t-1)

3 Calculate background using frame difference

FDBG(x,y,t)=$\frac{1}{3}\sum_{i=red,green,blue} FDGB_I$ (x,y,t)

4 Calculate background using median approximation

AMBD(x,y,t)=

$$\begin{cases} \mathrm{AMBD}(x,y,t-1)+1, & \mathrm{AMBD}(x,y,t-1) < f(x,y,t) \\ \mathrm{AMBD}(x,y,t), & \mathrm{AMBD}(x,y,t-1) = f(x,y,t) \\ \mathrm{AMBD}(x,y,t)-1, & \mathrm{AMBD}(x,y,t-1) > f(x,y,t) \end{cases}$$

The value of $\propto$ controls the effect of previous background to determine the new background.

II.3 Segmentation

Segmentation of the crowded video is done in the following 4 steps:

1. Background subtraction.
2. Noise removal
3. Thresholding
4. Post processing.

II.3.1 Background Subtraction

To detect the moving crowd each frame under process is subtracted by the estimated background image. For foreground detection we use simple algorithm given below, the algorithm is fast and gives good result.

ALGO. 2 Foreground detection

For each $f(x,y,t)$

Do

1. Calculate foreground_1 as
   Foreground_1 (x,y,t) = $f(x,y,t)$-FDBG(x, y, t)
2. Calculate foreground_2 as
   Foreground_2(x, y, t) = f(x, y, t) - AMBD(x, y, t)
3. Foreground is generate by taking average of Foreground_1 and Foreground_2
   Foreground(x, y,t)=$\frac{\text{foreground}\ 1+\text{foreground}\ 2}{2}$
4. Apply the median filter to reduce the noise.

II.3.2 Noise Removal

Theoretically the non-zero regions in the subtraction result are the location of the moving crowd. However, there are many other things presented in the practice. These unwanted things are called the background changes as a whole. {} classify the additional things except the moving crowd into three classes. They are the:

1. Noise,
2. Light changes and
3. The shadow of the moving object.

The gaussian noise presented in the image sequence captured by a camera. This noise is distributed on an image in a random manner. Therefore this type of noise cannot be removed by the background subtraction method. Many methods are applicable to reduce the Gaussian noise in an image. I use a simple method to reduce the noise in the foreground image generated in the previous step. The foreground image is convolved with an average mask which takes a local average computation for each pixel. The Gaussian noise can be effective reduced by this simple process and the changing regions appear [5].

The second class of the background changes is the light change. Light is the energy source for images. The charge-coupled device of a camera senses the light photons reflected from the objects in the image scene. The light changes will reflect on the image. Typical light sources are the sunlight and the indoor lamps. The image sequence obtained from an indoor camera may be influenced by the lamps and the sun light shining through the windows. The lamps may be turned on or off between two consecutive image frames [6]. Thus, it will result in a background change. The sunlight may shine on some local regions through the windows. These regions may move slowly with time.

The third class of the background changes is shadow of moving object. Shadow has the properties that it may appear or disappear at any time, it is always accompanied with a real motion object, and it again makes only intensity change. Based on the above observations ,the false detections can be eliminated by checking each feature vector's history.

Algo.3 Binary image conversion

1. For Foreground(x,y,t) determine the value of Foreground(x-1,y-1,t),Foreground(x-1,y,t),Foreground(x-1,y+1,t),Foreground(x,y-1,t),Foreground(x,y,t), Foreground(x,y+1,t),Foreground(x+1,y-1,t), Foreground(x+1,y,t),Foreground(x+1,y+1,t).
2. If all the above value is greater than Th

    Then

        BW(x,y,t)=1

    Else

        BW(x,y,t)=0

II.3.3 Post Processing

The initial binary image generated by segmentation process consists several closed regions. Each regions consists of a group of pixels. There are also some small noise regions because of irregular object motion and camera noise. These noisy regions are removed by simple application of 2-dimensional median filtering [7]. Morphological close and opening operation is adopted to connect the separated part and fill small holes inside the foreground caused by background subtraction.

III. Results and Analysis

For the implementation part MATLAB 2016b. is used. It stands for 'Matrix Laboratory'. MATLAB 2016b is a high level language for technical computing. It integrates computation, visualization and programming in an easy-to-use environment where problem and solutions are expressed in familiar mathematical notation. Typical uses include: MATLAB is an interactive system whose basic data element is an array that does not require dimensioning. This allows to solve many technical computing problems, especially those with matrix and vector formulation, in a fraction of the time it would take to write a program in a scalar non interactive language such as '*C*' and 'FORTRAN'.

MATLAB 2016b has evolved over a period of years with input from many users. In university environments, it is standard instructional tool for introductory and advanced courses in mathematics, engineering and science. In industry, MATLAB 2016b is the tool of choice for high productivity research, development and analysis. MATLAB 2016b features a family of add-on application-specific solutions called ‘toolboxes’. Very important to most users of MATLAB 2016b, toolboxes allow you to learn and apply specialized technology. Toolboxes are comprehensive collections of MATLAB2016b functions (M-files) that extend the MATLAB 2016b environment to solve particular classes of problems [8].

IV. Flow Chart of The proposed work

Crowd Localization is a very complex problem in computer vision. Attempts have been made to create a perfect crowd tracker. Many of the existing algorithms work under specific conditions, but all algorithms fail when it comes to wide-area surveillance videos where the area to be captured by video extends to square miles. The crowds in these videos appear blurred and feature extraction becomes extremely difficult. The problem becomes more complex when the effects of external environmental factors like smog, fog, haze, etc. are also considered. The presence of shadows and occlusions make tracking the crowd even more difficult.

The main aim of this research will be to develop fast algorithms that are able to address all the aforementioned challenges and also to achieve real-time performance. The following are the objectives of this research: To design an algorithm that can describe the crowd that has been selected by the user. This has to be implemented by taking into consideration the real-time constraints of the application as well as the accuracy of the tracking mechanism.

To implement a tracking mechanism that can handle occlusions [9]. To build an adaptive learning mechanism into the system, so the system learns about the crowd (its features as well as motion characteristics) in order to perfect the tracking mechanism.

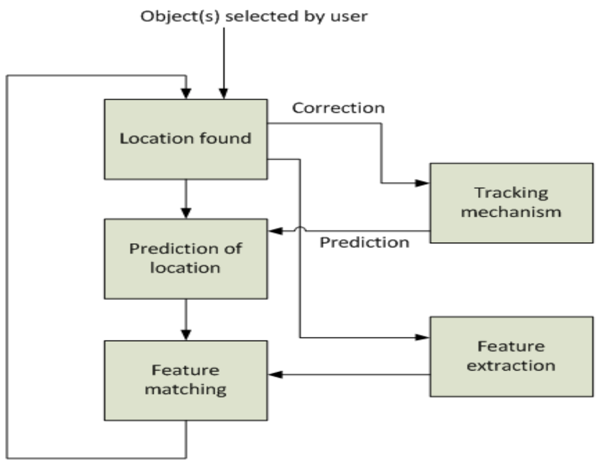

Figure2: Proposed Process of Crowd Detection and Tracking

V. Conclusion & Future Scope

Vision is one of the most heavily used sensors by humans. We interact with our environment mainly using visual sensors. Out of many high-level tasks that we can accomplish with vision, object detection stands out. Every day we recognize a multitude of familiar and novel objects. Many of the high-level tasks we perform are seamlessly integrated with object detection and in-fact we don't realize them. Object detection helps us to do a wide range of daily activities like moving around, interacting with people, reading, playing, etc. For better understanding, these tasks can be divided into those which involve implicit and explicit object detection.

Tracking crowd is a well-studied problem in computer vision, robotics, and related areas. The goal is to spatially and temporally localize each moving pedestrian in a video and compute its trajectory [12-15]. As autonomous robots and driverless cars are increasingly used in the physical world with tens or hundreds of pedestrians, it is important to track, and also predict motion and behavior at realtime rates. We also need real-time crowd tracking capabilities for surveillance activities The problem of crowd tracking has been extensively studied and a variety of techniques have been proposed. In many ways, pedestrians correspond to the most difficult categories of object tracking. Pedestrians tend to change their speed to avoid collisions with obstacles and other pedestrians. Large variations in their appearance and illumination make it hard for color-based template tracking algorithms to continuously track a pedestrian. In crowded scenes [11], the pair wise interactions between pedestrians can increase significantly and add to the complexity of predictive tracking schemes.

In this work a new method for crowd detection is presented which is based on background subtraction methods. For background estimation a new algorithm proposed which is very efficient and requires very less resource in terms of memory and time. The algorithm can generate the background image with low computational complexity and high efficiency. To determine the automatic threshold value SIFT algo. is used. For converting the gray scale image into binary image a new algorithm is proposed, which is based on the 8-connected neighborhoods.

I have also shown that the estimation of the crowd is better in comparison to the frame differencing method and approximated median method. The error in crowd estimation of the three methods is also compared and it is shown that the error in proposed method. There are still low in comparison to the other two methods.

References:

[1] Yang, D. B., &Guibas, L. J. (2003, October). Counting people in crowds with a real-time network of simple image sensors. In null (p. 122). IEEE.

[2] Alahi, A., Jacques, L., Boursier, Y., &Vandergheynst, P. (2009, December). Sparsity-driven people localization algorithm: Evaluation in crowded scenes environments. In Performance Evaluation of Tracking and Surveillance (PETS-Winter), 2009 Twelfth IEEE International Workshop on (pp. 1-8). IEEE.

[3] Havasi, L., &Szlßvik, Z. (2009, September). Using location and motion statistics for the localization of moving objects in multiple camera surveillance videos. In Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on (pp. 1275-1281). IEEE.

[4] Ge, W., & Collins, R. T. (2010, September). Crowd detection with a multiview sampler. In European Conference on Computer Vision (pp. 324-337). Springer, Berlin, Heidelberg.

[5] Alahi, A., Jacques, L., Boursier, Y., &Vandergheynst, P. (2011). Sparsity driven people localization with a heterogeneous network of cameras. Journal of Mathematical Imaging and Vision, 41(1-2), 39-58.

[6] Lo, K. H., & Chuang, J. H. (2011, September). Vanishing point-based line sampling for efficient axis-based people localization. In Image Processing (ICIP), 2011 18th IEEE International Conference on (pp. 521-524). IEEE.

[7] Blanke, U., Troster, G., Franke, T., &Lukowicz, P. (2014, April). Capturing crowd dynamics at large scale events using participatory gps-localization. In Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2014 IEEE Ninth International Conference on (pp. 1-7). IEEE.

[8] Chen, C. Y., & Shao, Y. (2015). Crowd escape behavior detection and localization based on divergent centers. IEEE Sensors Journal, 15(4), 2431-2439.

[9] Sabek, I., Youssef, M., &Vasilakos, A. V. (2015). ACE: An accurate and efficient multi-entity device-free WLAN

localization system. IEEE transactions on mobile computing, 14(2), 261-273.

[10] Chang, S. P., Chien, J. T., Wang, F. E., Yang, S. D., Chen, H. T., & Sun, M. (2016, October). Extracting Driving Behavior: Global Metric Localization from Dashcam Videos in the Wild. In European Conference on Computer Vision (pp. 136-148). Springer, Cham.

[11] Minaeian, S., Liu, J., & Son, Y. J. (2016). Vision-based target detection and localization via a team of cooperative UAV and UGVs. IEEE Transactions on systems, man, and cybernetics: systems, 46(7), 1005-1016.

[12] Yoo, J., Kim, H. J., & Johansson, K. H. (2016, October). Mapless indoor localization by trajectory learning from a crowd. In Indoor Positioning and Indoor Navigation (IPIN), 2016 International Conference on (pp. 1-7). IEEE.

[13] Zhou, S., Shen, W., Zeng, D., Fang, M., Wei, Y., & Zhang, Z. (2016). Spatialûtemporal convolutional neural networks for anomaly detection and localization in crowded scenes. Signal Processing: Image Communication, 47, 358-368.

[14] Alahi, A., Wilson, J., Fei-Fei, L., &Savarese, S. (2017, May). Unsupervised camera localization in crowded spaces. In Robotics and Automation (ICRA), 2017 IEEE International Conference on (pp. 2666-2673). IEEE.

[15] Chaker, R., Al Aghbari, Z., &Junejo, I. N. (2017). Social network model for crowd anomaly detection and localization. Pattern Recognition, 61, 266-281.